Genetic Assessment of Risk for Heart Disease in Northern Native Americans

Proposed Plan of Study, 15 Aug 96 by Jeffrey H. Simonson

Coronary heart disease (CHD) is the primary cause of death in Western societies. Elevated plasma cholesterol concentration is a major factor in atherosclerosis. Cholesterol is carried from its source of synthesis to its target in peripheral tissues by apolipoproteins. Most serum cholesterol is associated with low density lipoproteins (LDL). Apolipoprotein B (apoB) is the major protein constituent of LDL and binds to the LDL receptor (LDLR) in peripheral tissues. Variants of both apoB and LDLR, therefore, can significantly affect cholesterol levels. Several known mutations in the genes encoding apolipoproteins disrupt this targeting and increase or decrease plasma cholesterol concentrations. The best known mutation is a missense mutation on amino acid 3500 of apolipoprotein B. An Arg-->Gln change inhibits ApoB binding to the LDLR and prevents clearing of plasma LDL. Genetic factors affe plasma cholesterol levels and therefore contribute to the risk for CHD.

Abstract

Genetic factors affect plasma cholesterol levels and therefore contribute to the risk for coronary heart disease (CHD). Apolipoproteins carry cholesterol from its source of synthesis to its target where these proteins bind to receptors for receptor-mediated endocytosis. Several known mutations in the genes encoding apolipoproteins disrupt this targeting and increase or decrease plasma cholesterol concentrations. We propose a genetic assessment for CHD in Northern Native Americans based on frequencies of apolipoprotein mutations affecting plasma lipid and cholesterol levels. Studies investigating these mutations in Circumarctic peoples are limited. They do, however, suggest these groups have few mutations and are at lower risk for CHD than non-natives. We have chosen two loci already correlated with plasma lipid and cholesterol levels. The apolipoprotein B (APOB) gene codes for the main cholesterol carrying protein. This protein is also the ligand for uptake of low density lipoprotein (LDL). A mutation correlated with high levels of LDL is a G-To-C point mutation in amino acid number 3500. This mutation causes hypercholesterolemia by preventing apoB from binding to its receptor. The second locus is the apolipoprotein AI/CIII/AIV gene cluster. Though the mechanism is unclear, studies show a correlation between several polymorphic sites and plasma lipid levels. We plan to sample from all three Northern Native American language phyla: Amerind, Na-Dene, and Eskimo-Aleut. We will collect hair samples and extract DNA from the root cells. Following amplification by the polymerase chain reaction, we will detect mutations by digesting with a restriction enzyme designed to cut only in the presence of the mutation. This study will give an estimate of the genetic risk for CHD in Circumarctic peoples. Since the groups normally show less heart disease than non-natives, we may be able to estimate the relative influences of environment and genetics on this major human disease.

Introduction

Coronary heart disease (CHD) is the primary cause of death in Western societies (Innerarity 1990). The link between elevated levels of plasma cholesterol and atherosclerosis, the major cause of CHD and stroke, is well established (Lusis 1988, Grundy 1990). Though many environmental factors contribute to CHD risk, genetic factors also contribute by affecting cholesterol concentration. Cholesterol is carried from its source of synthesis to peripheral target cells by lipid-carriers called apolipoproteins (Lusis 1988, Young 1990). Most serum cholesterol is associated with low density lipoproteins (LDL). Apolipoprotein B (apoB) is the major constituent of LDL and binds to the LDL receptor (LDLR) in peripheral tissues. Variants of both apoB and LDLR, therefore, can significantly affect cholesterol levels. Apolipoprotein A (apoA) is coded for by a linked, multigene family and serves as the ligand for receptor-mediated endocytosis of cholesterol in high density lipoproteins (HDL) (Lusis 1988, Hixson 1991). These loci exhibit considerable genetic variation (Kessling 1992) and investigators have identified several apolipoprotein mutations causing increased or decreased cholesterol levels (Hobbs 1987, Cole 1989, Young 1989, Ludwig 1990, DeBenedictis 1993, Miserez 1993, Kamboh 1996). Here I propose an assessment of risk for CHD in Northern Native Americans based on frequencies of apolipoprotein mutations affecting plasma lipid and cholesterol concentrations. Populations having high frequencies of the disease-causing mutations will be at high risk for CHD; the converse is also true. The Evenki (Kamboh 1996) and the Dogrib (Cole 1989) both have low risk for CHD because they have few of the cholesterol-increasing mutations. Since Circumarctic peoples exhibit reduced levels of genetic variation when compared to Amerind groups distributed farther south (Shields 1993), they likely have low risk for CHD. Therefore, I propose to test the hypothesis that Northern Native Americans have a lower risk for heart disease than other populations.

Methods

I plan to test this hypothesis by direct assessment of apolipoprotein allele frequencies. All loci I have chosen are thought to affect plasma lipid and cholesterol levels. The APOB gene and the apolipoprotein A-I/C-III/A-IV gene cluster are good candidates for CHD risk assessment (Cole 1989, Soria 1989, Kamboh 1996, Lusis 1988). Apolipoprotein B ApoB serves as the ligand for uptake of LDL by the LDL receptor (Blackhart 1986). Two forms of apoB exist; apoB-100, a large protein associated with transport of plasma LDL, and apoB-48, a post-transcriptionally shortened version of apoB-100. ApoB-48 mediates absorption of fats from the intestine (Young 1989). Mutations in the ligand region of the APOB gene disable binding to the LDLR and cause a change in cholesterol level because cells cannot take in LDL by receptor-mediated endocytosis (Innerarity 1990). The allele showing the best correlation with high LDL levels is the 3500 mutation in apoB-100 (Soria 1989). The 3500th amino acid is changed from arginine to glutamine by a G-to-C point mutation in the ligand region. This mutation directly causes hypercholesterolemia in both homo- and heterozygotes because plasma apoB cannot bind to the LDLR (Blackhart 1986, Soria 1989, Ludwig 1990). I will assay this mutation by digesting polymerase chain reaction (PCR) products with the MspI restriction enzyme as in Soria (1989). Young (1989) found another point mutation causing a change in cholesterol level. A T-to-C mutation at nucleotide 6381 in APOB-48 changes an arginine codon to a premature stop codon and disables translation of the ligand region necessary for LDLR binding. This mutation, however, causes low plasma cholesterol levels because intestinal epithelial cells cannot take up LDL. I will assay this mutation by digesting PCR products with the TaqI restriction enzyme as in Young (1989). Apolipoprotein AI/CIII/AIV Gene Cluster Mutations in this linked gene cluster have been shown to affect plasma lipid levels (Lusis 1988, Cole 1989, Kamboh 1996). I have chosen several polymorphic sites defined by the restriction sites XmnI, PstI, SstI, and PvuII described by Cole et al. (1989) in their study of the Dogrib. These alleles are involved in determining lipid levels, though the mechanism is unclear. Cole (1989) showed the PstI site in the APOAI gene is rare in the Dogrib; the XmnI site, due to a 300 base pair (bp) deletion, is also absent. Cole (1989) additionally showed allele frequencies in the Dogrib vary significantly from other populations and are not in Hardy-Weinberg equilibrium. I also want to look at the Fnu4HI restriction site of the APOA4 gene. Kamboh (1996) found this unique marker for white populations completely absent in the Evenki; these Siberians have low serum lipid levels. Sample Populations I intend to sample populations from all three Northern Native American language phyla as defined by Greenberg (1986): Amerind, Na-Dené, and Eskimo-Aleut. Blood samples from Eastern Siberian Yupiks (Eskimo-Aleut) are already available. Tlingits (Na-Dené) from Southeast Alaska and Inupiaqs (Eskimo-Aleut) from Northern Alaska will be new to this study. These populations are valuable because pedigrees already exist for investigation of the mode of inheritance of apolipoprotein alleles. Molecular Methods The protocols I will use for these molecular analyses are already well established. I plan to use PCR to amplify regions surrounding the particular mutation using primers described for each locus. PCR is performed on chelex-extracted DNA (Walsh 1991) from either blood or hair. These products are then cut with restriction enzymes or sequenced. For sequence analysis, purified PCR products are cycle-sequenced with the Sanger (1977) dideoxy method using flourescently labeled dinucleotides. The purified sequence products are then visualized on the Applied Biosystems 373A automated DNA sequencer located in the Institute of Arctic Biology's Core Facility for DNA Analysis. I plan to sequence a representative sample of all apolipoprotein PCR products to assay for mutations other than ones elucidated by the restriction fragment length polymorphic (RFLP) analyze. RFLP study involves cleaving PCR products with the appropriate restriction enzyme and visualizing them on either polyacrylamide or agarose gels stained with ethidium bromide. I will then determine allele frequencies by analyzing the size and numbers of each band. To support the link between these mutations and their effects on cholesterol levels, I will sample serum cholesterol concentrations of the Siberian Yupiks. Since these are valuable samples and contain low volumes, the sample size must be minimized. The Fairbanks Memorial Hospital can determine serum cholesterol concentration using as little as 50uL of blood. I plan to get a better picture of the relationships between the native groups by using my existing set of mitochondrial DNA (mtDNA) sequence. I have sequenced the mitochondrial control region of 30 Navajos and 93 Siberian Yupiks. This will help estimate the risk for CHD in related populations because mtDNA can elucidate relationships not discernible by nuclear genes (Wilson 1985). Data Analysis Risk for coronary heart disease will be estimated by calculating allele frequencies for entire populations based on sample genetic markers. I will analyze these allele frequencies using the Hardy-Weinberg theorem and other techniques for estimating selection and drift (Golding 1994). I will also define relationships between separate populations by generating phylogenetic trees. These phylogenies are built using the software programs PAUP (Swofford 1989) and Phylip (Felsenstein 1989). Combining my mtDNA sequences for Siberians and Navajos with the apolipoprotein data will help make these trees more accurate. This analysis may then be used to estimate relationships between other populations not studied. In this way, my study of several Northern Native American populations can be used to assess the risk for heart disease in other closely related peoples. Expected Results Based on work with the Dogrib (Cole 1989) and Evenki (Kamboh 1996) populations, I expect my study populations will differ in apolipoprotein allele frequencies from other populations. My mtDNA sequence data should also show this divergence. These populations are not likely to have the APOB-100 3500 mutation; their chance of having the APOB-48 mutation is unclear. I expect Siberian Yupiks to show fixation at the Fnu4HI site of APOA4 as in Kamboh (1996). The four haplotypes for the apolipoprotein A-I/C-III/A-IV gene cluster described by Cole (1989), and used in this study, should depart from Hardy-Weinberg expectations just as in the Dogrib. The Dogrib are Amerindians, so Navajos will likely exhibit similar frequencies. Furthermore, Shields et al. (1993) showed a close relation between Northern Na-Dené and Eskimo-Aleuts; this leads me to predict a close relationship between Tlingits and Inupiaqs even though they belong to different language phyla (Greenberg 1986). Generally, Circumarctic peoples should show low risk for CHD as in Cole (1989) and Kamboh (1996). These data may then be useful for determining a cause for the higher incidence of CHD in non-native populations.

Literature Cited

Anderson S, et al. 1981. Sequence and organization of the human genome. Nature 290:457-465. Blackhart BD, et al. 1986. Structure of the human apolipoprotein B gene. J. Biol. Chem. 261(33):15364-15367.

Cole SA, Szathmary EJE, and Ferrell RE. 1989. Gene and gene-product variation in the apolipoprotein A-I/C-III/A-IV cluster in the Dogrib Indians of the Northwest Territories. Am. J. Hum. Genet. 44(6):835-843.

DeBenedictis G, et al. 1993. RFLPs for the APOB gene: Comparative study between Greeks and southern Italian peoples. Hum. Biol. 65(3):401-411.

Felsenstein J. 1989. Phylogenetic Inference Package, Version 3.2 (Univ. Washington Seattle).

Golding B. 1994. Non-neutral evolution: Theories and molecular data.

Chapman & Hall, NY. Greenberg JH, Turner CG, Zegura SL. 1986. The settlement of the Americas: A comparison of the linguistic, dental and genetic evidence. Curr. Anthropol. 27:477-497.

Grundy SM. 1990. Cholesterol and coronary heart disease. JAMA 264(23):3053-3059.

Hixson JE and Powers PK. 1991. Restriction isotyping of human apolipoprotein A-IV: Rapid typing of known isoforms and detection of a new isoform that deletes a conserved repeat. J. Lipid Res. 32:1529-1535.

Innerarity TL, et al. 1990. Familial defective apolipoprotein B-100: A mutation of apolipoprotein B that causes hypercholesterolemia. J. Lipid Res. 31:1337-1349.

Kamboh HI, et al. 1996. Population distributions of APOE, APOH, and APOA4 polymorphisms and their relationships with quantitative plasma lipid levels among the Evenki herders of Siberia. Hum. Biol. 68(2):231-243.

Kessling A, et al. 1992. Patterns of association between genetic variability in apolipoprotein (apo) B, apo AI-CIII-AIV, and cholesterol ester transfer protein gene regions and quantitative variation in lipid and lipoprotein traits: Influence of gender and exogenous hormones. Am. J. Hum. Genet. 50(1):92-106.

Ludwig EH, McCarthy BJ. 1990. Haplotype analysis of the human apolipoprotein B mutation associated with familial defective apolipoprotein B100. Am. J. Hum. Genet. 47(4):712-720.

Lusis AJ. 1988. Genetic factors affecting blood lipoproteins: The candidate gene approach. J. Lipid Res. 29:397-429.

Miserez AR, et al. 1993. Polymorphic haplotypes and recombination rates at the LDL receptor gene locus in subjects with and without familial hypercholesterolemia who are from different populations. Am. J. Hum. Genet. 52(4):808-826.

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.

Shields GF, Schmiechen AM, Frazier BL, Redd A, Voevoda MI, Reed JK, Ward RH. 1993. mtDNA sequences suggest a recent evolutionary divergence for Beringian and Northern North American populations. Am. J. Hum. Genet. 53(3):549-562.

Soria, LF, et al. 1989. Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. Proc. Natl. Acad. Sci. USA 86: 587-591.

Swofford DL. 1989. PAUP, Phylogenetic Analysis Using Parsimony, Version X (Illinois Natural History Survey, Champaign, IL). Walsh SP, Metzger DA, Higuchi R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques 10(4):506-513.

Young SG, et al. 1989. Familial hypobetalipoproteinemia associated with a mutant species of apolipoprotein B (B-46). N. Engl. J. Med. 320(24):1604-1610. Young SG. 1990. Recent progress in understanding apolipoprotein B. Circulation 82(5):1574-1594.

Wilson AC, et al. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linnaean Soc. 26:375-400.